Synergistic multi-wavelength remote sensing versus a posteriori combination of retrieved products: Application for the retrieval of atmospheric profiles using MetOp-A

F. Aires, O. Aznay, Christophe Prigent, M. Paul, Frédéric Bernardo

Synergistic multi-wavelength remote sensing versus a posteriori combination of retrieved products: Application for the retrieval of atmospheric profiles using MetOp-A

F. Aires,[1,2,3] O. Aznay,[4] C. Prigent,[3] M. Paul,[1] and F. Bernardo[2]

Received 21 November 2011; revised 25 July 2012; accepted 31 July 2012; published 21 September 2012.

[1] In this paper, synergy refers to a process where the use of multiple satellite observations makes the retrieval more precise than the best individual retrieval. Two general strategies can be used in order to use multi-wavelength observations in an inversion scheme. First, the multi-wavelength observations are merged in the input of the retrieval scheme. This means that the various satellite observations are used simultaneously and that their possible interactions can be exploited by the retrieval scheme. Second, each multi-wavelength observations are used independently to retrieve a same geophysical variable and then, these independent retrievals are combined a posteriori using for example a simple weighted averaging. In this paper, it is shown that the first approach provides better synergy results: The retrieval is better suited to optimize the use of all the information available because they are provided to the algorithm simultaneously. In particular, the retrieval process is able, in this case, to exploit the possible interactions between the various input information. The two retrieval approaches are tested and compared using an application for the retrieval of atmospheric profiles and integrated column quantities (temperature, water vapor, and ozone) using MetOp-A observations from IASI, AMSU-A and MHS instruments. Although real satellite observations are considered in this analysis, the results are dependent on the correlation structure in the training data set (i.e. ECMWF analysis) used to calibrate the retrieval algorithm. However, it can be seen that the infrared and microwave observations have a good synergy for the retrieval of atmospheric temperature, water vapor, and for ozone thanks to an indirect synergy.

Citation: Aires, F., O. Aznay, C. Prigent, M. Paul, and F. Bernardo (2012), Synergistic multi-wavelength remote sensing versus a posteriori combination of retrieved products: Application for the retrieval of atmospheric profiles using MetOp-A, *J. Geophys. Res.*, *117*, D18304, doi:10.1029/2011JD017188.

1. Introduction

[2] Synergy refers to a process where the use of multiple satellite observations makes the retrieval more precise than the best individual retrieval. In *Aires* [2011], it has been shown that various synergy mechanisms exist: (1) additive synergy, the simpler mechanism, where the addition of multiple information on a same geophysical variable increases naturally the retrieval accuracy; (2) indirect synergy, where the relationships between the geophysical variables are exploited by the retrieval scheme; (3) the non-linear synergy that acts when interaction terms of the satellite observations are relevant for the retrieval; and (4) de-noising synergy that refers to situations where the instrument noise of the observations are correlated. By exploiting all these synergy mechanisms, it is possible to benefit from the simultaneous use of all observations, obtaining better performances than the best individual retrieval.

[3] Synergy has been exploited in various applications. For example in *Worden et al.* [2007] and *Landgraf and Hasekamp* [2007], infrared and UV observations are used simultaneously to retrieve atmospheric ozone. Infrared and near-infrared are also combined to retrieve CO_2 [*Christi and Stephens*, 2004], CH_4 [*Razavi et al.*, 2009] or CO [*Worden et al.*, 2010]. In *Aires et al.* [2011b], it was shown that the NN inversion model is particularly well adapted to exploit the synergy among satellite observations: The synergy of InfraRed (IR) and MicroWave (MW) observations from MetOp-A was tested for the retrieval of atmospheric temperature and water vapor profiles. These

[1]Laboratoire de Météorologie Dynamique/IPSL/CNRS, Université de Paris VI/Jussieu, Paris, France.
[2]Estellus, Paris, France.
[3]Laboratoire de l'Etude du Rayonnement et de la Matière en Astrophysique, CNRS, Observatoire de Paris, Paris, France.
[4]Institut Pierre-Simon-Laplace, Université de Paris VI/Jussieu, Paris, France.

Corresponding author: F. Aires, CNRS/IPSL/Laboratoire de Meteorologie Dynamique, Université Pierre et Marie Curie, case 99, 4, place Jussieu, FR-75252 Paris CEDEX 05, France. (filipe.aires@lmd.jussieu.fr)

0148-0227/12/2011JD017188

tests were performed using theoretical simulations from Radiative Transfer (RT).

[4] In this paper, real observations from IASI (Infrared Atmospheric Sounding Interferometer), AMSU-A (Advanced Microwave Sounding Unit-A) and MHS (Microwave Humidity Sounder) instruments will be used instead of the RT simulations. Furthermore, the retrieval scheme will retrieve atmospheric profiles not only for temperature and water vapor, but also for ozone. The total column amount of water vapor and ozone will be retrieved as well.

[5] Two general strategies can be used in order to benefit from multi-wavelength observations in an inversion scheme. First, the multi-wavelength observations are merged as inputs to the retrieval scheme. This means that the various satellite observations are used together and simultaneously by the retrieval scheme. Second, each wavelength observation is used independently, to retrieve a given geophysical variable and then, these independent retrievals are combined a posteriori using for example a simple weighted average. These two synergy strategies will be tested to identify the best approach.

[6] First, the data sets used in this study will be described in section 2, along with the necessary pre-processing of the data. In section 3, a preliminary analysis of the satellite observation information content will be conducted. The retrieval methodologies and synergy measures will be presented in section 4. The results will be described in section 5 using real observations; the retrieval of both integrated quantities and profiles will be assessed. Finally, section 6 will conclude this study and present the perspectives for this work.

2. Data Sets

2.1. The Satellite Observations

[7] MetOp-A is Europe's first polar-orbiting satellite dedicated to operational meteorology. MetOp is a series of three satellites to be launched sequentially over 14 years, forming the space segment of EUMETSAT's Polar System (EPS). The first satellite of the series was launched on October 19, 2006. The use of MetOp-A instruments in the operational centers is beneficial for the retrieval of atmospheric temperature, humidity and ozone profiles, wind speed and direction, and trace gases. MetOp-A flies in a polar morning orbit with a 9:30 AM/PM local solar time Equator crossing time.

[8] AMSU-A is a cross-track radiometer scanning between $\pm$ 48.3° from nadir, with a total of 30 Earth fields-of-view of 3.3° per scan line. The swath width is approximately 2000 km. It provides a nominal spatial resolution of 48 km at nadir. The instrument completes one scan every 8 s. AMSU-A is divided into two separate modules: (1) AMSU Module A-1 with twelve sounding channels in the 55 GHz O_2 band [*Mo*, 1996] and one at the 89 GHz window (channels 3 to 15); (2) AMSU Module A-2 with channels 1 and 2 at 23.8 and 31.4 GHz. AMSU-A is used in conjunction with the High-resolution Infrared Radiometer Sounder (HIRS) instrument to estimate the global atmospheric temperature and humidity profiles from the surface to the upper stratosphere, i.e. $\simeq$50 km [*EUMETSAT*, 2009]. AMSU-A measurements also provide precipitation and surface information including snow cover, sea-ice concentration, and soil moisture.

[9] The MHS sounder is designed to measure the atmospheric water vapor profile. It possesses three channels in the H_2O line at 183.31 GHz plus two window channels at 89 and 150 GHz [*Hewison and Saunders*, 1996]. MHS scans the Earth from left to right, in a vertical plane. Each swath is made up of 90 contiguous individual pixels sampled every 2.67 s. The scan is also synchronized with the AMSU-A1 and A2 instruments.

[10] IASI is a state-of-the-art Fourier transform spectrometer based on a Michelson interferometer coupled to an integrated imaging system. It was developed by the French space agency CNES [*Chalon et al.*, 2001]. The optical interferometry possesses 8461 channels, measuring the infrared radiation emitted from the Earth at a fine spectral resolution (0.25 cm^{-1} unapodized) in the range of 3.62 μm (2760 cm^{-1}) and 15.5 μm (645 cm^{-1}). This enables the instrument to retrieve temperature and water vapor profiles in the troposphere and the lower stratosphere. IASI also measures radiances that are sensitive to the concentrations of ozone, carbon monoxide, methane and other compounds [*Coheur et al.*, 2005; *Razavi et al.*, 2009; *Pommier et al.*, 2010; *George et al.*, 2009; *Maddy et al.*, 2009]. For optimum operation, the IASI measurement cycle is synchronized with that of the AMSU-A. This instrument was designed to reach accuracies of 1 K in temperature and 10% in water vapor with vertical resolutions of 1 km and 2 km respectively for cloud-free scenes. The Instantaneous Field Of View (IFOV) of IASI is 12 km at Nadir, but four IASI pixels are averaged and projected into one MW pixels, so the nominal spatial resolution is here 48 km. Total 8461 channels are divided into three bands, with (645–1210 cm^{-1}) the first band, (1210–2000 cm^{-1}) the second band and (2000–2760 cm^{-1}) the third band. The channels of this third band have a signal-to-noise too low (measured radiance is low, and instrument noise is too high) with no real additional information for temperature, water vapor and ozone compared to the first two bands. As a consequence, only the first 5421 channels of the first two bands will be used in this study. It should be noted that band 3 can however be exploited for other applications such as the retrieval of carbon monoxide [*Pommier et al.*, 2010; *George et al.*, 2009; *Maddy et al.*, 2009].

[11] The volume of data is considerable for the remaining 5421 channels of IASI. In order to sample as well as possible the seasonal variability while keeping the volume of data reasonable, four weeks of observations have been gathered: The first weeks of January, April, July, and October 2008. In order to limit practical difficulties (space memory, computing time and sea-ice mask), only pixels from $\pm$30° in latitude have been kept in the experiments.

[12] It has been shown in *Aires et al.* [2002a, 2011b] that it is very important, before applying the retrieval algorithm, to pre-process the IASI data using a Principal Component Analysis (PCA). This is useful to compress the data, and also to reduce the instrument noise. In order to process the satellite observations in a systematic way, a PCA is used for the two wavelengths considered (MW and IR). Based on the percentage of variance explained by each PCA component, twenty components have been used for the IR, and twelve for the MW (please note that the number of components is highly dependent on the application under study). This is an

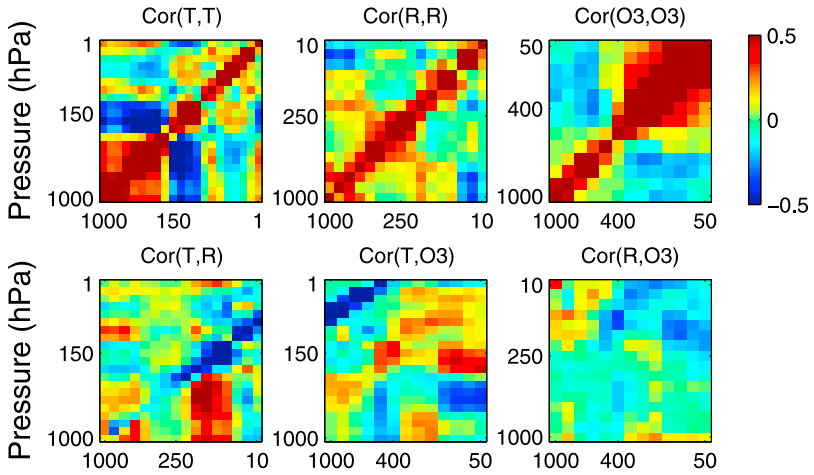

Figure 1. (top) Auto-correlation matrices for atmospheric profiles of (from left to right) Temperature (T), Relative humidity (R), and Ozone (O_3). (bottom) Correlations among the atmospheric profiles (from left to right): T (vertical axis) & R, T (vertical axis) & O_3 and R (vertical axis) & O_3.

estimate of the number of exploitable independent pieces of information that can be extracted from each wavelength.

[13] In this study, the observations from AMSU-A, MHS and IASI onboard MetaOp-A will be used.

2.2. The ECMWF Analyses

[14] The atmospheric profiles and surface properties from the 6-hourly operational global analyses from the Integrated Forecasting System (IFS) of the European Centre for Medium Range Forecasting (ECMWF) [*Uppala et al.*, 2005] are used in this study. These analyses are provided on a 1.125° × 1.125° regular grid. The following information is kept: The temperature, water vapor (relative humidity in % hereafter) and ozone profiles, on 43 pressure levels ranging from 1000 to 1 hPa. These levels have been interpolated from the initial twenty-one levels in order to be used with the RTTOV code. The cloud cover information from the ECMWF analysis is also kept, to filter out the cloudy scenes in the retrieval process. The radiances from MetOp platform that will be used for the retrieval are actually assimilated in the ECMWF analysis using the same radiative transfer code, RTTOV. This is a good thing because this means that the physical link that exists between the real observations and the ECMWF analysis is respected in the way our data set is built (which is to put in space/time coincidence the real observations and the ECMWF analysis). This link is at the basis of the learning of the retrieval schemes that will be designed in this study.

[15] The atmospheric profiles are provided on 43 pressure levels. However, there are no real ozone observations for the higher-level layers. The same holds for water vapor in to upper layers. As a consequence, these variables have been suppressed in the retrieval scheme. The selection that has been chosen corresponds to: (1) All the levels are kept for the temperature retrievals; (2) the five highest atmospheric layers are suppressed for the water vapor (pressure lower than 10 hPa); (3) and the eight highest layers are suppressed for the retrieval of ozone (lower than 50 hPa). Ozone is present in these higher layers, but unfortunately, the ECMWF analysis has not enough information on it, the retrieval scheme cannot be trained in this data set. An alternative would be to use an independent estimation of the ozone (from GOME II instrument for example) but then, the coherency between the temperature, water vapor and ozone profiles would be lost. The improvement of the ozone data set will be considered in forthcoming studies.

[16] Figure 1 represents the correlations between the geophysical atmospheric profiles, among themselves (Figure 1, top) and between them (Figure 1, bottom). It can be seen that the vertical correlations are stronger for temperature than for water vapor. The ozone has a strong vertical correlation in the upper troposphere where the ozone content is higher. The correlations between the profiles indicate that the temperature profile is strongly correlated to the water vapor for pressures lower than 250 hPa. There exists also a strong correlation between the temperature and ozone for both surface and upper layers. The correlation between water vapor and ozone is less important but it is still significant [*Vaughan et al.*, 2005]. Note that through the correlations between the geophysical variables, a satellite observation not physically sensitive to a particular geophysical variable, can still be statistically related, indirectly, to it. Therefore, this satellite observation can provide valuable information that can be exploited by the retrieval scheme.

2.3. The Aerosols

[17] The aerosol contamination would degrade our temperature, water vapor and ozone retrievals especially for IR observations. In order to filter out situations with too high Aerosol Optical Thicknesses (AOTs), an independent source of information on aerosols needs to be used. A climatology of aerosols in 492 in situ locations has been obtained from AERONET stations from 1996 to 2010. AERONET is a globally distributed network of automated ground-based instruments and data archive system, developed to support the aerosol community. The instruments used are Cimel spectral radiometers that measure the spectral extinction of the direct Sun radiance [*Holben et al.*, 1998]. The AOTs are determined using the Beer-Bouguer Law in several spectral bands. For this study, level-2 data are used and consist of AOTs at 440 nm, 675 nm and 870 nm retrieved, at least, every 15 minute during daytime. Level-2 data are cloud-free

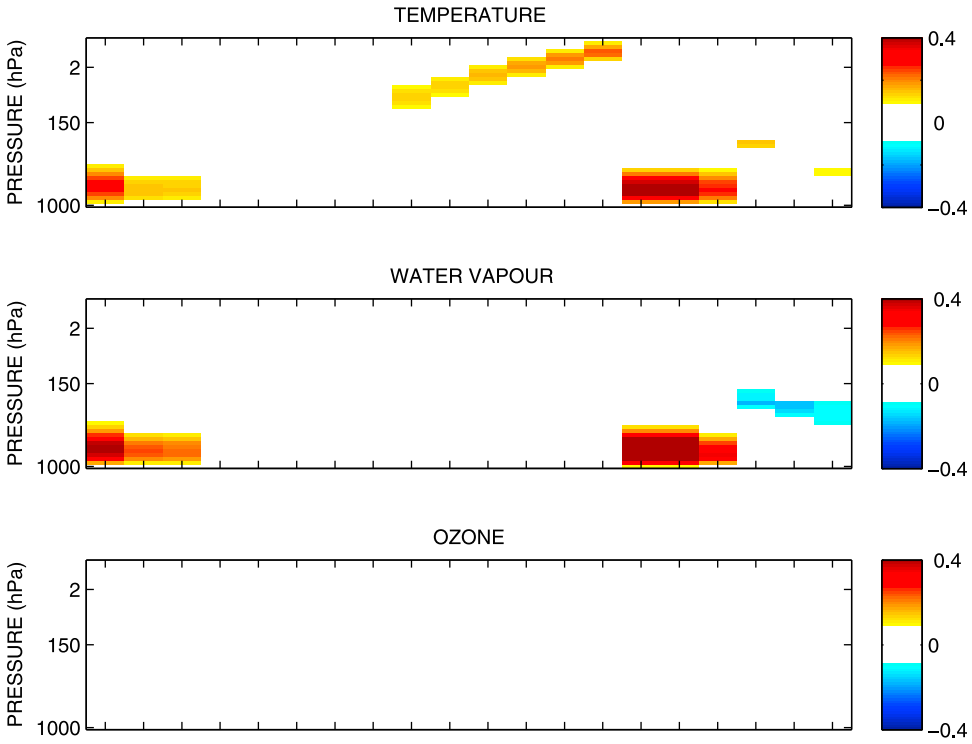

Figure 2. Jacobian of the AMSU-A + MHS observations with respect to (top) temperature, (middle) water vapor, and (bottom) ozone atmospheric profiles for a typical ocean scene over the Tropics.

and quality assured retrieved from pre- and post-field calibrated measurements [*Smirnov et al.*, 2000]. The estimated accuracy in the AERONET AOTs is between ±0.01 and ±0.02 depending on the wavelength, for an air mass equal to one [*Dubovik et al.*, 2000].

[18] A spatial interpolation scheme has been applied to obtain AOT fields at a horizontal resolution of 1.125° × 1.125° (compatible with the ECMWF analyses). This spatial interpolation scheme is rather crude (a bilinear interpolation). In order to obtain better AOT fields, a dedicated interpolation scheme would have to be developed specifically for AOT fields. This could be based, for example, on a PCA of complete AOT fields. But since such a data set is not available, this approach cannot be used here. The alternative would be to develop an AOT retrieval scheme based on the satellite observations, but this is beyond the scope of this study.

2.4. Spatiotemporal Coincidences

[19] The satellite observations need to be matched in space and time. The two instruments being on-board the same satellite, good coincidences are obtained. The collocation of the MW observations is easy, the instruments have been designed to facilitate this step: Each AMSU-A pixel is associated to the corresponding 3 × 3 higher resolution pixels of MHS. The IR observations from IASI are then projected into the MW pixels using a "closest pixel" rule. The maximum scanning angles of IASI, AMSU and MHS are respectively 48.2°, 48.33° and 49.44°. The final resolution of the data set is the MW resolution (i.e., 48 × 48 km at nadir).

[20] The analysis and aerosol data sets (sections 2.2 and 2.3) are then projected into the satellite observations. A time threshold of 30 minute is tolerated for this collocation so satellite orbits are only kept around the analysis time steps at 0, 6, 12 and 18 h UTC. For each satellite pixel, the coincidence is performed using the closest grid point of the analysis. This means that there can be multiple use of the same analysis grid cell.

[21] Only oceanic situations are kept. The total cloud cover from the ECMWF is adopted to reject all the cloudy situations.

3. Preliminary Analysis

3.1. Sensitivity Analysis

[22] The Jacobian of the RTM are estimated for the three instruments considered here, namely AMSU-A, MHS and IASI. The Jacobians are estimated using RTTOV simulations on perturbed input profiles but analytical Jacobians could have been used too. The perturbations are chosen to be 1 K for temperature and 10 % for humidity and ozone (in ppmv). The readers are referred to *Garand et al.* [2001] for an inter-comparison study of such Jacobians.

[23] Figure 2 represents the temperature, humidity and ozone Jacobians for AMSU-A, MHS for a typical tropical situation over the ocean. As expected, the MHS instrument is more sensitive to changes in humidity than to changes in temperature. AMSU-A provides temperature information in the tropopause and for pressures higher than 100 hPa. Its window channels are sensitive to both temperature and water vapor. As expected, the MHS instrument is more sensitive to water vapor. There is no sensitivity to ozone in the MW observations. Figure 3 represents similar Jacobians but for the IASI instrument. It can be seen that temperature retrieval can be based on the 645–770 cm^{-1} CO_2 absorption band (strong sensitivity to temperature and Jacobians well distributed along

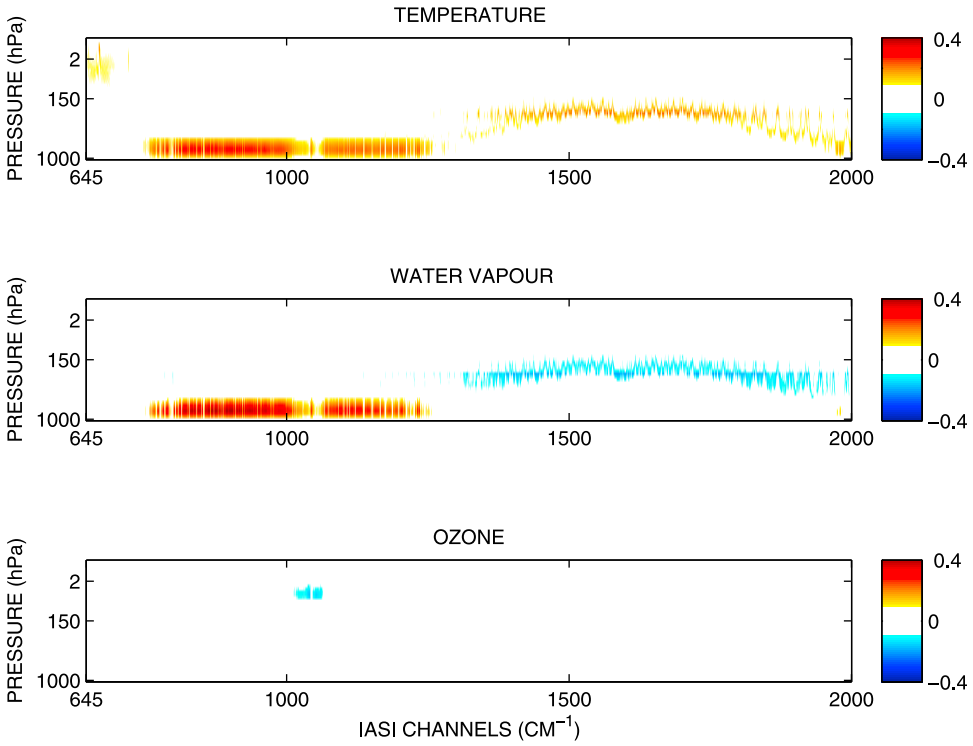

Figure 3. Jacobian of the IASI observations with respect to (top) temperature, (middle) water vapor, and (bottom) ozone atmospheric profiles for a typical ocean scene over the Tropics for the first two bands of IASI.

the atmospheric column). The water vapor Jacobian follows closely the temperature Jacobian, both in terms of pressure levels and half-height width. Channels located in the 1210–1650 cm^{-1} water vapor absorption band are sensitive to temperature and water vapor between 700 and 160 hPa, it is then necessary to isolate in the retrieval process, the signal of these two variables. Channels located between 980 and 1160 cm^{-1} shows the sensitivity to O_3 that allows for the retrieval of O_3. The magnitudes of Jacobians are comparable for the MW and IR wavelengths in terms of temperature and water vapor. However, IASI provides a higher vertical resolution for water vapor in the upper troposphere, and for temperature in the whole troposphere. Furthermore, the high-dimension of IASI observations have consequences for the retrievals, both negative (computation time for the retrieval and necessity to perform a dimension reduction on the observed spectra) and positive (such as the redundancy of information in the channels that can be exploited to reduce instrument noise [*Aires et al.*, 2002a]).

3.2. Correlation Matrix Analysis

[24] The correlation matrices between the temperature, water vapor and ozone ECMWF atmospheric profiles with respect to the collocated satellite observations in the MW, respectively IR, are represented in Figure 4, respectively Figure 5. Only blue and red colors should be considered in the analysis of these correlation matrices, they represent the strong negative and positive correlations. These correlation matrices using real observations provide a good and simple information content analysis. Not only can they be related to the sensitivity of the observations to the geophysical parameter of interest, but also to the correlation of the observations with a parameter that is itself correlated with another parameter of interest (indirect synergy).

[25] For instance, let us examine the correlation between the microwave observations and the water vapor. The major water vapor line is located at 183 GHz, with a much weaker line at 22 GHz. Outside these two lines, the contribution of the water vapor continuum increases with the frequencies. One would expect the correlation between the observations and the water vapor to be higher (in absolute value) in the absorption lines, and to increase with the continuum absorption. For pressures below 200 hPa, the amount of water vapor is limited and correlation structures between the water vapor and the observations in the H_2O lines or in window channels are very likely related to the correlation of the water vapor between the different levels in the ECMWF analysis. From the analysis of the Jacobian in water vapor, it was shown that the sensitivity to water vapor was high in the 183 GHz channel but very low in the O_2 band at 57 GHz. The correlation with the observations in the O_2 band is to be attributed, not to the sensitivity of the *TB* to the water vapor directly, but to the correlation between the water vapor with the temperature. Note that the correlation with the temperature and with the water vapor shows rather similar patterns (top and middle figures), in relation with the correlation between water vapor and temperature. By the same token, the microwave observations are not expected to be sensitive to the ozone variation. The correlation between the microwave observations and the ozone in the O_2 band is actually due to the correlation between temperature and ozone, in the ECMWF reanalysis (Figure 1). The fact that the correlations are not identical to the physical sensitivities does not

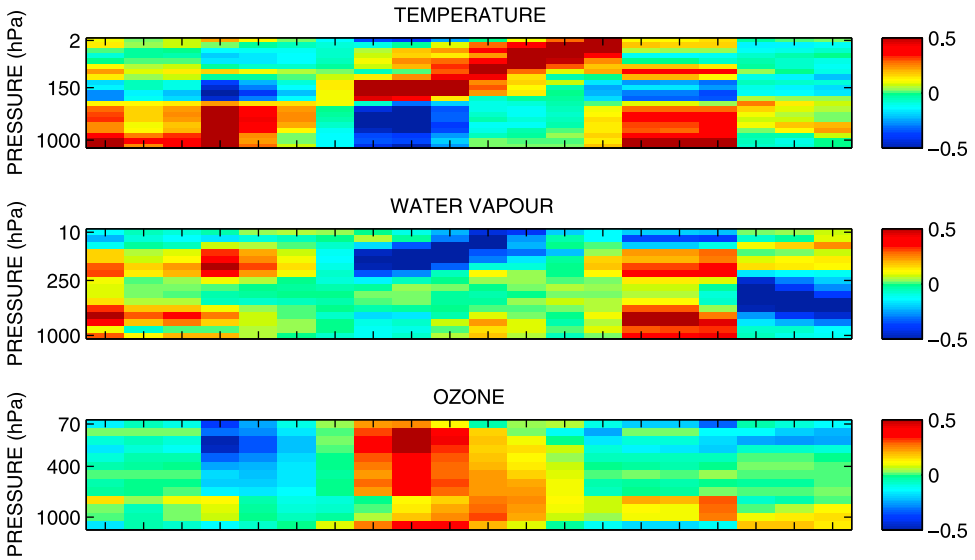

Figure 4. Correlations between AMSU-A and MHS observations and the ECMWF atmospheric profiles of (top) temperature, (middle) water vapor, and (bottom) ozone.

mean that they are spurious results. They are real correlations and it is legitimate exploiting them in the retrieval scheme.

[26] Similar conclusions can be derived for the infrared. The strong water vapor band above 1500 cm^{-1} induces large correlation between the observations and the water vapor. This frequency band is also very correlated to the temperature, as expected from the analysis of the Jacobians in temperature. In the CO_2 absorption band, the correlation between the IASI observations at about 645 cm^{-1} and the temperature is large, negative in the lower atmosphere and positive in the higher atmosphere, related to the inversion in temperature in the atmosphere. There is a gap in the correlation near 300 hPa for both temperature and water vapor, suggesting that the retrieval of these layers will be difficult. We also observe large correlation in this frequency range with the water vapor and the ozone, due to the intrinsic correlation between the three atmospheric variables in the

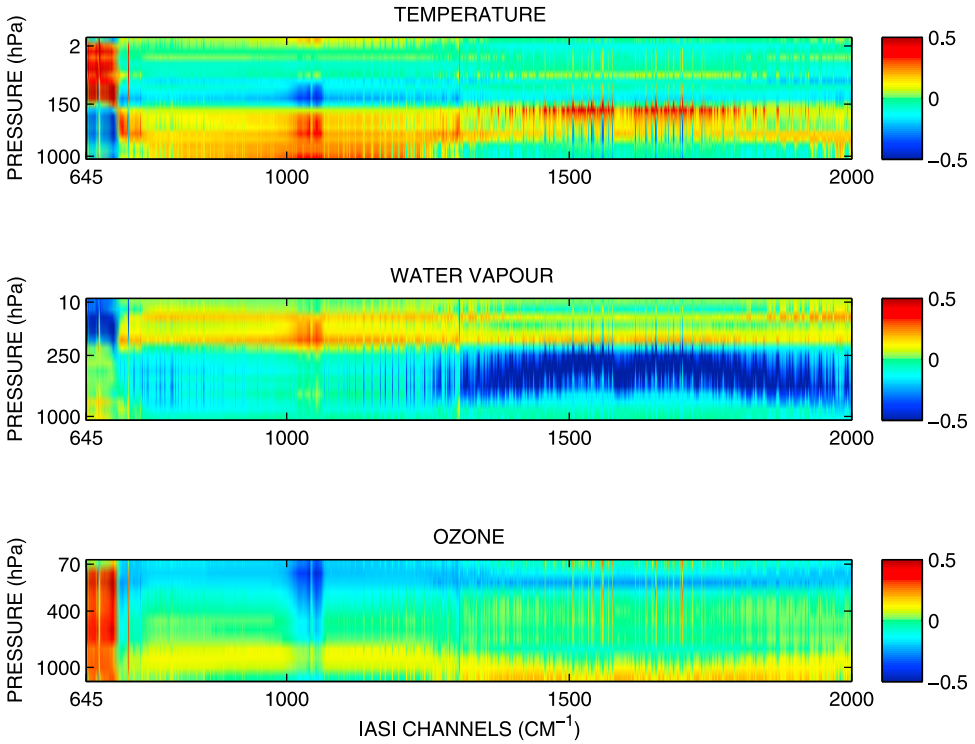

Figure 5. Correlations between IASI observations and the ECMWF atmospheric profiles of (top) temperature, (middle) water vapor, and (bottom) ozone.

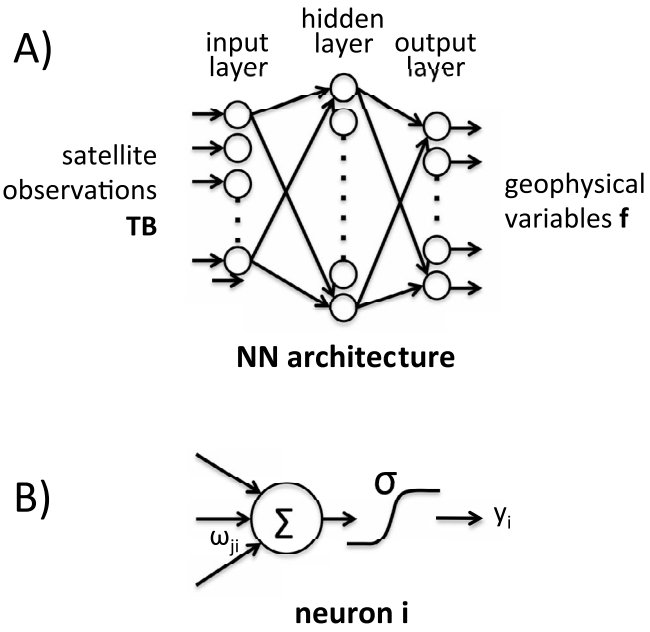

Figure 6. (a) Architecture of a one hidden layer neural network and (b) an individual processing neuron.

atmospheric profiles. The O_3 band 1000–1070 cm^{-1} induces similar effect on the water vapor and temperature.

3.3. Jacobian Versus Statistical Information

[27] The retrieval algorithms that will be used in this study are neural networks trained on a database of real satellite observations and coincident ECMWF atmospheric profiles. Like most retrieval schemes, the NN approach not only exploits the direct link between the observations and selected atmospheric variables (analyzed in section 3.1), but also the correlations between the different variables at the different levels (Figure 1). As for most statistical techniques, the link between the observations and selected atmospheric variables is described by some statistical link, not by the radiative transfer Jacobians.

[28] The simultaneous exploitation of direct physical sensitivities and indirect statistical correlations can be observed in most retrieval schemes. For example, let consider a 1D-var retrieval scheme:

$$x_a = x_b + \left(B^{-1} + H^t R^{-1} H\right)^{-1} H^t R^{-1} (y_0 - \mathcal{H}(x_b)),$$

where x_a is the retrieval (analysis), x_b is the background state, B is the covariance matrix of the background state, H is the Jacobian of the radiative transfer, R is the covariance matrix of the observation and the radiative transfer errors, y_0 is the satellite observations, and $\mathcal{H}$ is the radiative transfer [*Kalnay et al.*, 1996]. This retrieval uses directly the Jacobians $\mathcal{H}$ but it also exploits some statistical constraints based on coherent relationships (e.g., dynamic or thermodynamic constraints) between the temperature, water vapor and ozone profiles. This is expressed by the background state x_b or by the error covariance matrix of the background state, B.

[29] We also have to be aware that limitation of the quality of the atmospheric profiles that are used to train a statistical retrieval (such as the one pointed out for ozone in section 2.2) strongly limits the quality of the retrieval of the corresponding variable.

4. Retrieval Methodology

4.1. Architecture of the Neural Network Inversion Models

[30] NN techniques have proved very successful in developing computationally efficient remote sensing algorithms. The Multi-Layered Perceptron (MLP) model [*Rumelhart et al.*, 1986] is selected here. It is a non-linear mapping model: Given an input *TB* (i.e., observed Brightness Temperature), it provides an output *f* (i.e., the geophysical variables to retrieve) in a non-linear way. In this paper, a NN model with only one hidden layer will be considered (Figure 6). Each layers in the NN is composed of individual neurons. A neuron performs first a weighted average of its inputs from the previous layer. The so-called synaptic weights are associated to each connection between two neurons. These weights represent the NN parameters to be defined during the training stage. The NN chosen in this study is a fully connected MLP (i.e., every neuron has a connection with all the neurons of the previous layer). Once this weighted average is done, a non-linear sigmoid function is applied. The final output of a neuron i is given by:

$$y_i = \sigma\left(\sum_{j=1}^{N} w_{ji} x_j\right),$$

where $(x_j; j = 1, \ldots, N)$ are the N inputs of the neuron, w_{ji} is the synaptic weight between neuron j and i, and σ is a sigmoid function [*Bishop*, 1996] (bias terms are also present in this model, but they are suppressed here for simplicity of the presentation).

[31] The MLP model is defined by the number of input neurons (i.e., the size of the inputs, number of channels), the number of outputs (i.e., the size of the geophysical variables to retrieve) and the number of neurons in the hidden layers that control the complexity of the model. A study has to be conducted to define the optimal number of neurons in the hidden layer. A balance needs to be found: Too many free parameters in the model can result in over-learning (over-parameterization) leading to degraded generalization properties. On the contrary, too few free parameters will yield under-parameterization and bias error of the model.

[32] The neural network used in this study can be represented by a very simple function: $f(TB) = W_2 \cdot \sigma(W_1 \cdot TB)$, where f represents the geophysical variables, TB represents the satellite observations and W_1 (resp. W_2) is the matrix of weights from the input to the hidden layer (resp. from the hidden to the output layer). The Jacobian of this function can be derived for any state TB: $\frac{\partial f(TB)}{\partial TB} = W_2 \cdot \sigma'(W_1 \cdot TB) \cdot W_1$, where σ' is the derivative of the sigmoid function σ. The state-dependency of this Jacobian results from the presence of the sigmoid function σ, otherwise the model would be linear and the Jacobians would remain constant for all states TB. This state-dependency is important for the exploitation of the synergy in satellite observations TB: The NN is able to adapt the sensitivity of its outputs to the inputs based on the state. It can use only one source of information when the others sources are not pertinent or it can combine them in a

Table 1. Architecture of the NN Retrieval Models

Configuration	Number of Inputs	Hidden Neurons	Number of Outputs
Temp-IASI	20	20	21
Temp-MW	12	20	21
Temp-IR+MW	32	25	21
WV-IASI	20	20	16
WV-MW	12	20	16
WV-IR+MW	32	25	16
O_3-IASI	20	20	13
O_3-MW	12	20	13
O_3-IR+MW	32	25	13

non-linear way when necessary. Complex interactions among the satellite observations can be exploited by the NN. In order to benefit from this "non-linear synergy", it is clearly necessary to present the various observations in the inputs of the NN. Combining independent retrievals a posteriori would not allow to exploit the complex interactions among the satellite data that the NN is able to perform. Readers are invited to read *Aires* [2011] for additional details.

[33] Since the NN performances will be compared using different configurations of inputs (only the IR, only the MW or both), it is important to address the relative stability of the training. A NN with more inputs will have more free parameters and more information to exploit; this can make difficult the NN training, and the learning step can become much longer. An over-parameterization can lead to over-training if the learning process is not regularized. Good testing and validation data sets, stopping criterion, and the multiple initializations of the weights for the learning tend to reduce this problem. Table 1 represents the number of neurons in the input, hidden and output layers for the various configuration that are used in this paper.

4.2. The Learning, Testing, and Validation Data Sets

[34] The NN is trained to reproduce the behavior described by a database of samples composed of inputs (i.e., the real observations TB) and their associated outputs (i.e., the geophysical variables f), for $e = 1,\ldots, N$ with N the number of samples in the training database. Provided that enough samples (TB^e, f^e) are available, any continuous relationship, as complex as it is, can be represented by a MLP [*Hornik et al.*, 1989]. Furthermore, in *Cybenko* [1989], a theorem shows that a two-hidden layer NN is able to represent any discontinuous function which is an important feature for inverse problems.

[35] A quality criterion that measures the discrepancies between the NN outputs and the desired targets from the learning data set has to be defined. In this study, the outputs of the NN are the atmospheric profiles (temperature, water vapor and ozone) and integrated column quantities of water vapor (i.e., $TCWV$) and ozone (i.e., TCO_3). In this paper, the weighted least squares criterion is used so that the same weight $\frac{1}{5}$ is given to the three profiles and the two integrated quantities. Furthermore, each retrieved atmospheric layer is associated to a weight of $\frac{1}{5} \times \frac{1}{Nb}$, where Nb is the number of layers in the profile. This quality criterion is minimized during the learning of the NN. This choice of weights is ad hoc. The absolute value of these weights is not significant but their relative values will affect the training process and the resulting NN statistics. The NN outputs with higher weights will be retrieved with more precision.

[36] The learning algorithm used to train the NN is the classical "back-propagation" algorithm. This optimization technique has long proved its efficiency for such problems.

[37] Over the whole data set of coincident real satellite observations and analyses, 60% are kept for the learning, 20% for the testing, and 20% for the validation. These three data sets have been selected by random selection. Since the number of samples in each one is very large (i.e. few thousands), their statistical characteristics are similar (law of large numbers). The training of the NN (i.e., the calibration of the retrieval scheme) is performed on the learning data set.

[38] The testing data set is used in parallel to the training process: at each learning step, the learning data set is used to improve the parameters of the NN model. The testing data set is not directly used to change the NN parameters, but it is also used in each step of the learning to monitor the results of the NN in an independent data set. This allows measuring the generalization capacity of the NN, i.e., its ability to perform retrievals on new data. During the learning, the generalization errors are monitored and the learning is stopped when, after a decrease, they start to increase. This procedure avoids the over-training of the retrieval, i.e., the problem of an algorithm that performs very well on its learning data set, but poorly on new data.

[39] The testing data set is used several times, first to measure the generalization capacities of the NN on each step of the learning process, second for all the tested NN configurations, and finally for all considered retrieval configurations (e.g., angle, optical thickness). As a consequence, the whole NN selection process could "learn" the testing data set (i.e., be biased toward it) and the evaluation of the generalization errors could become misleading. To avoid this problem, another independent data set is used: The validation data set. It is only used to estimate the retrieval errors on an independent data set, once the learning is done, and once the model is chosen.

4.3. Regime Selection

[40] The retrieval algorithm has been developed for various configurations: (1) Viewing Zenith Angle (VZA) equal to 0, 10, 20, 30 or 40°; (2) Solar Zenith Angle equal to (SZA) 40, 50 or 60°; (3) Aerosol Optical Thickness lower than 0.05 (see section 2.3); (4) only over oceanic surfaces; and (5) only for clear situations (see section 2.2). This regime selection results in $5 \times 3 = 15$ data sets designed to train 15 specialized retrieval schemes. It should be pointed out that the cloud flag from the ECMWF analysis is not ideal, but an a posteriori radiative transfer test is done in order to measure the discrepancy of ($TB_{sim} - TB_{obs}$) in order to flag incoherent situations. Furthermore, in order to limit the impact of the aerosols in the visible measurements, only AOT <0.05 have been kept (see section 2.3).

4.4. Physical Versus Empirical Retrievals

[41] In order to train the NN, two strategies could be used: First, the inversion can be trained on a learning database composed of the atmospheric profiles from the ECMWF analysis, along with the simulated observations in IR and MW bands derived from RT calculations (instead of real

observations). This type of inversion is said to be a "physical" inversion, as it uses a physical RT model. However, this procedure requires a preliminary calibration step to insure that the RT simulations and the real observations are compatible (i.e., have similar statistics).

[42] Second, the training can be performed on a learning database composed of the satellite observations and collocated profiles from ECMWF analysis. This type of scheme is said to be an "empirical" inversion because no RT model is used to solve the inverse problem. As commented earlier, the fact that the real radiances from MetOp-A are assimilated in the analysis is not a problem, on the contrary, it strengthen the link between the radiances and the analysis. However, the analyses are not perfect and some incoherencies between the radiances and the atmospheric situations in the analysis can be present in the data set for multiple reasons: bad cloud flag, presence of aerosols, wrong profiles, etc. These discrepancies translate in a weaker link between radiances and profiles, which implies a reduced quality of the retrieval. Another inconvenient of the empirical approach is that the use of new observations from another satellite would require the retraining of the full retrieval scheme, where the physical approach would not require the modification of the inversion step, only the calibration would need to be redone.

[43] The first approach involves explicitly two operations, namely the calibration of the satellite data and the actual retrieval [*Aires et al.*, 2010]. The second approach involves only one transformation of the real observations: It mixes the calibration and the retrieval in a unique procedure. Both methods could lead to satisfactory results, they both have advantages and disadvantages. However, the second one is preferred here because obtaining a good calibration procedure for both the IR and MW is a particularly difficult task.

[44] The use of the ECMWF analyses as the "truth" for the training of the NNs is justified because these analyses are reliable for temperature and water vapor profiles. All the available satellite observations as well as all the in situ measurements from the weather stations have been assimilated in one of the best weather/climate models. The ozone data quality from the analysis system is probably not so high, but it is important here to use a coherent data set of atmospheric profiles. This means that the present results on ozone retrievals are probably less reliable.

4.5. Synergy Measures

[45] Information theory allows measurement of the information carried out by a variable or an ensemble of variables [*Shannon*, 1948]. This type of measures (e.g. mutual information or entropy) can be used to quantify the synergy. For example, the synergy of two variables v_1 and v_2 can be defined by:

$$syn(v_1, v_2) = I(v_1, v_2) - (I(v_1) + I(v_2)),$$

the difference between the information I conveyed by the pair and the information conveyed by the two events independently. Information I can be measured for example using the entropy [*Brenner et al.*, 2000]. This approach allows computing how much information is conveyed by the patterns, but it does not indicate what particular message these patterns convey. This type of measure allows constructing a quantitative measure for the significance of compound patterns, this is and essential first step in understanding the synergy among satellite observations. However, an additional measure of the information should take into account the transformation between the inputs and the outputs of the retrieval scheme.

4.5.1. Synergy Factor

[46] A synergetic scheme refers to an algorithm that uses simultaneously or hierarchically the observations of two or more spectral ranges in order to obtain a more accurate retrieval than the independent retrievals. We define a synergy factor of a retrieval scheme using R sources of information ($x_1,\ldots, x_R$) (each one can be multivariate) as the ratio of the errors of the retrieval using the best single information, $Min_{i=1,\ldots,R}E(x_i)$, with the errors of the retrieval using all the sources of information, $E(x_1,\ldots, x_R)$. In terms of percentage of synergy, this corresponds to:

$$F_{syn} = 100 \cdot \left(\frac{Min_{i=1,\ldots,R}E(x_i)}{E(x_1,\ldots,x_R)} - 1 \right). \quad (1)$$

There is synergy when this quantity is positive. This synergy measure can be used for any type of retrieval algorithm that uses multiple wavelength observations, including the a posteriori combination of products that is described in the following section. In this study, the errors are estimated in the validation data set (see section 4.2).

4.5.2. A Posteriori Combination Versus Synergetic Data Fusion

[47] In this study, the goal is to merge synergistically the IR and MW observations directly as inputs to the retrieval algorithm. This data fusion or merging approach is unusual. In many of the retrieval schemes that use multiple source of observations, independent retrievals from each instrument observation are combined a posteriori instead of been merged as inputs of the retrieval. An example of such a posteriori combination is given, for example, in *Liu et al.* [2011] where two soil moisture estimates from two different algorithms using passive and active microwave observations are simply averaged. However, this a posteriori combination by simple averaging can be optimized.

[48] Let $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ a vector including n related observations of the same geophysical variable x such that:

$$Y = K \cdot x + \varepsilon \quad (2)$$

where K is the n-dimensional unit vector and uncertainties ε follow a Gaussian distribution $\mathcal{N}(0, S_\varepsilon)$. Covariance matrix S_ε can be diagonal if the estimates in Y have uncorrelated errors, but if there is a way to estimate a full covariance matrix, then it can be used directly in the following.

[49] It is interesting to decrease the uncertainty on the estimation of x by combining its multiple estimates in Y. For example, the estimates can be combined linearly: $\hat{x} = A \cdot Y$. The objective is then to optimize vector A of the linear model in order to decrease uncertainties in $\hat{x}$. A least-squares criterion or a maximum a posteriori optimization principle can be used in order to obtain A [*Rodgers*, 1990]:

$$A = \left(K^t S_\varepsilon^{-1} K\right)^{-1} K^t S_\varepsilon^{-1} \quad (3)$$

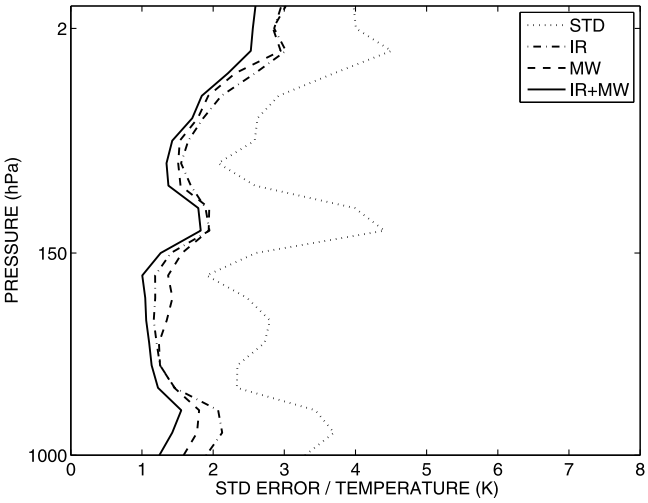

Figure 7. RMS errors for the temperature retrievals from IR (dot-dashed), MW (dashed), IR+MW (continuous) instruments. The natural STD of the temperature is also provided (in dotted line) for comparison purpose.

and the uncertainties are estimated with

$$Q = \left(K^t S_\varepsilon{}^{-1} K\right)^{-1}. \quad (4)$$

This is what is done in optimal interpolation or assimilation.

[50] For example, if S_ε is diagonal, the linear a posteriori combination becomes a simple weighted average based on the uncertainties of the two retrievals. Combining two independent retrievals f_{IR} and f_{MW} using IR and MW observations, each one with uncertainty estimates σ_{IR} and σ_{IR}, is optimal when using:

$$\hat{f} = \frac{\sigma_{MW}^2}{\sigma_{IR}^2 + \sigma_{MW}^2} \cdot f_{IR} + \frac{\sigma_{IR}^2}{\sigma_{IR}^2 + \sigma_{MW}^2} \cdot f_{MW}. \quad (5)$$

The theoretical uncertainty related to this estimator is given by:

$$\hat{\sigma} = \sqrt{\left(\frac{\sigma_{MW}^2}{\sigma_{IR}^2 + \sigma_{MW}^2}\right)^2 \cdot \sigma_{IR}^2 + \left(\frac{\sigma_{IR}^2}{\sigma_{IR}^2 + \sigma_{MW}^2}\right)^2 \cdot \sigma_{MW}^2}. \quad (6)$$

In the following, the "simple averaging" is defined as the a posteriori combination with all weights equal to one, the "weighted averaging" is defined as the a posteriori combination when the weights of each variable is based on their uncertainty (equation (6)), and the "full-covariance matrix averaging" is used when the full covariance matrix of uncertainties is taken into account (equation (4)).

[51] Other more sophisticated a posteriori combinations could also be considered, in particular, regime-dependent combinations that would take into consideration the state-dependency of the individual retrieval uncertainties. Histogram matching techniques have also been popular in some communities.

[52] In the following, the results of the retrieval scheme that performs synergetic merging of the satellite observation in the inputs of the retrieval scheme will be compared to the results of the independent retrievals using equation (5), i.e. a posteriori combination. The same amount of data is used in both of these approaches so it is very interesting to compare the results, in order to see if the merging of the satellite data can exploit the potential interactions among them.

5. Retrieval Results

[53] The retrieval of the atmospheric profiles is first presented and the synergy factors are presented for each atmospheric layer. The retrieval of the integrated quantities ($TCWV$ and TCO_3) is then analyzed, in particular, investigating the a posteriori combination.

[54] The retrieval statistics are provided for the following configuration: The aerosol content is low (between 0 and 0.05), the viewing zenith angle = 40° (the higher the angle, the more difficult the retrieval becomes), but the solar zenith angle is composited for the various configurations (see section 4.3). Note that similar results have been obtained for the other configurations. The tests are performed over ocean under clear sky conditions. Statistics are calculated using the validation data set (i.e., data not used in the learning or generalization data sets), see section 4.2.

5.1. Temperature, Water Vapor and Ozone Profiles

[55] The statistics are provided for retrievals using only MW, only IR, or IR+MW observations. Figure 7 shows the RMS errors and variability of the retrieval of the atmospheric temperature profile. It is interesting to note that the MW retrieval is better than the IR, especially in the lower troposphere, except in some layers around 300–100 hPa. The IR plus MW data fusion of the NN retrieval benefits from a strong synergy. Its RMS statistics range from 1 to 1.5 K in most atmospheric layers. The synergy estimations for these temperature profile retrievals will be analyzed in the following. The results obtained in this section are based on a large data set of situations so the differences between the various configurations are statistically significant.

[56] Similar statistical results are given in Figure 8 for the retrieval of the atmospheric water vapor profile. The IR

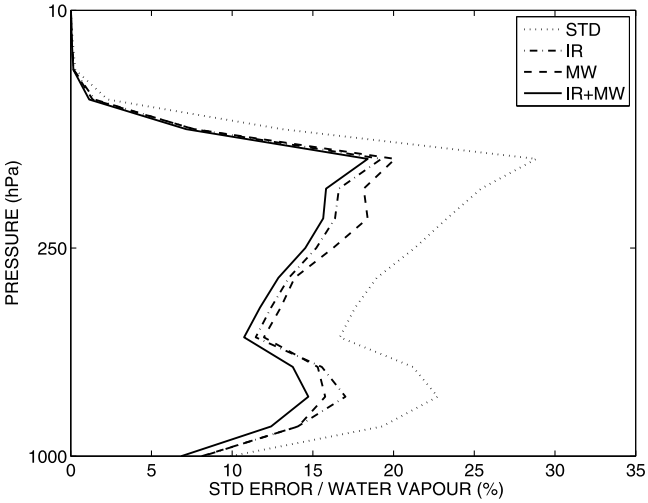

Figure 8. RMS errors for the water vapor retrievals from IR (dot-dashed), MW (dashed), IR+MW (continuous line) instruments. The natural STD of the temperature is also provided (in dotted line) for comparison purpose.

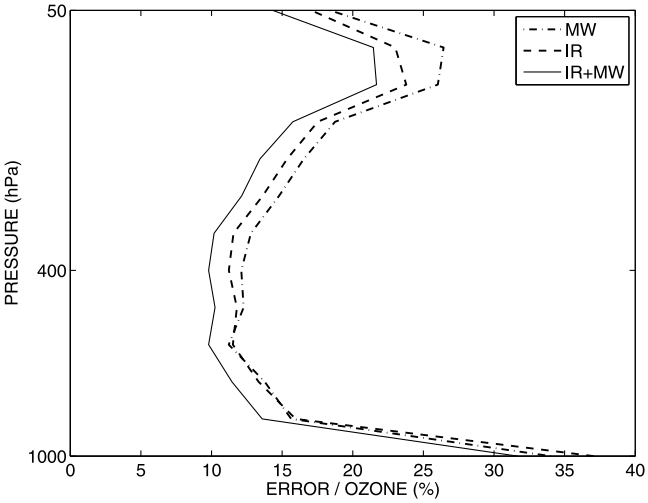

Figure 9. Mean absolute errors, in percentage, for the ozone retrievals from MW (dot-dashed), IR (dashed), and IR+MW (continuous line) observations.

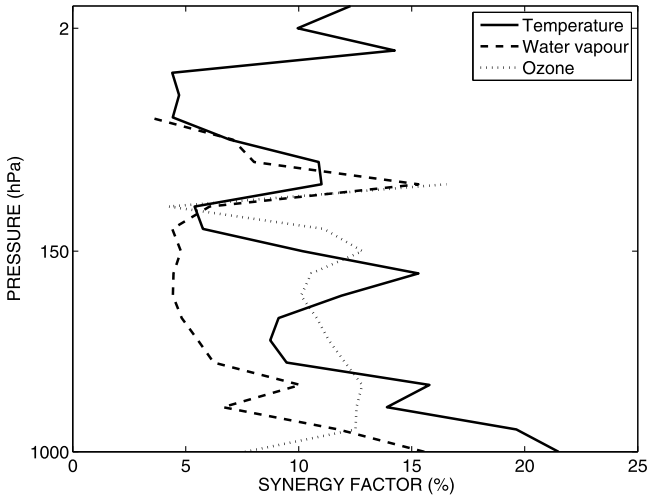

Figure 10. Synergy statistics for the retrieval of the temperature (continuous), water vapor (dashed) and ozone (dotted) atmospheric profiles. The retrieval uses the IR, MW observations from AMSU-A, MHS and IASI instruments.

observations from IASI appear to be more informative than the MW measurements from MHS/AMSU-A. This is particularly true for the upper troposphere above 300 hPa. The IR+MW data fusion has retrieval uncertainties statistics always lower than the best independent retrieval. The improvement is higher than one percent, which represents a synergy factor of six to ten percent.

[57] Since the ozone profiles have very different ranges of variability in the lower or higher troposphere, it is more convenient to represent the standard error of the retrieval in terms of percentage of errors. It can be seen in Figure 9 that errors are higher at high altitudes around 100 hPa (i.e., where the ozone content is higher) and near the surface (i.e., where the ozone content is very low, less than $0.2 \cdot 10^{-6}$ Kg/Kg). Except for few lower atmospheric layers, the IR from IASI provides the best information for the ozone profile retrieval. The synergy is very good when IR and MW are merged. The overall retrieval of the ozone atmospheric profile has a good quality with error levels lower than 30% and often lower than 15%. It should be noted, again, that our absolute uncertainty might be underestimated because the ozone profiles from the ECMWF analyses might be too simple compared to real ozone profiles. Furthermore, this study focuses on $\pm 30°$ in latitude, with possible limitations of the ozone variability in this region (note that the limitation of the range of variability of the variable to estimate is also a difficulty for the retrieval scheme). It could be surprising that MW observations help improving the IR retrieval of ozone since it has been shown in the Jacobian study of section 3.1 that the MW measurements are not sensitive to ozone. However, it has been seen in section 3.2 that the MW brightness temperatures are correlated to the ozone profile (Figure 4). This illustrates well that a 1D-var inversion scheme [*Kalnay et al.*, 1996], or any retrieval scheme that is based on the use of the Jacobians [*Bowman et al.*, 2006], does not use the same information than a statistical retrieval scheme, such as the NN, that exploits the correlations.

[58] Figure 10 represents the synergy measured when IR and MW observations are combined to retrieve temperature, water vapor and ozone atmospheric profiles. This synergy factor is based on equation (1). Synergy factor ranges from 10 to 15% for temperature, for all the considered atmospheric layers. The synergy for the ozone retrieval has similar characteristics. The synergy for the *WV* is overall very significant, between 5 and 15%, in agreement with the retrieval of *TCWV* in section 5.2. Synergy for water vapor is also positive for most atmospheric layers, except for the two top layers where it is known that low information content is provided by the satellite observations. Numerical instabilities can also be present.

5.2. The Integrated Quantities

5.2.1. Evaluation of the Neural Network Retrieval

[59] In this section, the retrieval of the *TCWV* and TCO_3 variables are performed using the data fusion/merging principle described in section 4: The NN model uses simultaneously the IR+MW observations to perform the retrieval.

[60] Table 2 shows some retrieval statistics for both *TCWV* and TCO_3, when for the IR, MW, and IR+MW configurations. This table provides the RMS errors of the retrievals, the correlation coefficient between retrieved and the desired variable, and the synergy factor when both IR and MW measurements are used in the NN retrieval. The MW observations are the best information for the *TCWV* (RMS = 3.3 $kg.m^{-2}$), better than the IR observations from IASI (RMS = 4.5 $kg.m^{-1}$). However, it should be noted that the

Table 2. RMS, Correlation and Synergy Statistics for the Retrieval of *TCWV* and TCO_3 Using IR, MW, and IR+MW Satellite Observations With the Neural Network Retrieval Scheme

	IR	MW	IR+MW
RMS error *TCWV* ($kg.m^{-2}$)	4.5	3.3	2.9
RMS error TCO_3 (DU)	6.6	8.2	4.9
Correlation *TCWV*	0.89	0.94	0.96
Correlation TCO_3	0.92	0.87	0.96
STD error TCO_3 (%)	2.6	3.2	1.9
Synergy *TCWV* (%)			11.7
Synergy TCO_3 (%)			26.0

Table 3. A Posteriori Combination Versus Data Fusion Retrieval Statistics for *TCWV* and TCO_3[a]

	A Posteriori Combination						Data Fusion
	Theoretical Average	Retrieval Average	Theoretical Weighted	Retrieval Weighted	Theoretical Full	Retrieval Full	Retrieval NN Total
RMS *TCWV* ($kg.m^{-2}$)	2.8	**3.6**	2.7	**3.4**	2.6	**3.3**	**2.9**
RMS TCO_3 (DU)	5.5	**6.3**	5.1	**6.1**	4.9	**6.0**	**4.9**
Syn. *TCWV* (%)		**−6.64**		**−1.88**		**0.12**	**11.7**
Syn. TCO_3 (%)		**4.35**		**6.60**		**9.09**	**26.0**

[a]For the a posteriori combination, both theoretical and real statistics are provided for the simple averaging, the weighted average and the full covariance matrix approaches. The data fusion principle uses the NN retrieval. Both approaches use IR and MW observations. Results for actual retrievals are represented in bold, contrarily to theoretical estimations.

cloud flag that was used to filter the cloudy situations in this study is based on the total cloud cover from the ECMWF analysis (see section 2.2) and this is not a perfect cloud flag information. Furthermore, the presence of clouds affect more the IR than the MW observations. Therefore, the cloud contamination could explain why MW observations seem to retrieve better the temperature profile. The MW observations are very important for the retrieval of the *TCWV* since they might be less affected by cloud contamination. The data fusion of the IR and MW in our retrieval scheme results in a very interesting synergy factor, 11.7%, with RMS = 2.9 $kg.m^{-2}$. The correlation statistics follow these RMS statistics. The *TCWV* retrieval in the IR+MW/SYN configuration has a 0.96 correlation with the target, which is a very good result.

[61] For the retrieval of the TCO_3, the best information is provided by the IR observations from the IASI instrument (RMS = 6.6 DU) followed by the MW (RMS = 8.2 DU). This could be surprising considering that MW observations are not physically sensitive to O_3. However, the IR and MW numbers are very close. Furthermore, MW observations and O_3 content have shown strong correlations in Figure 4. MW appeared to be linked to O_3 (correlation = 0.87), through an indirect correlation: The MW is related to temperature and water vapor, and the temperature and water vapor are themselves related to O_3. Our processing of the IR observations is not perfect (see sections 2.2 and 2.3): The cloud and aerosol flags are not optimal and this is particularly a problem for the IR. The synergy obtained when IR and MW are used is important, equal to 26%, with a RMS down to 4.9 DU. The synergy operates even better for the TCO_3 than for the *TCWV*. The two instruments have a significant contribution when used together. The standard error for the retrieval of ozone is also provided in percentage, as this is a common way to measure the quality of the ozone retrieval.

5.2.2. Satellite Data Fusion Versus a Posteriori Combination of Products

[62] In this section, a investigation is conducted to determine whether a posteriori combinations of products provides similar results compared to the satellite data merging. Three a posteriori combinations will be considered: The simpler one where the two independent estimates from the IR and MW are averaged, the weighted average that takes into account the uncertainty of each retrieval, and the full covariance approach that utilizes the correlation of errors (see section 4.5.2).

[63] It is possible to estimate the theoretical retrieval uncertainties from these a posteriori combinations by using the individual uncertainties of the two independent retrievals (see equation (5)). These theoretical estimates are compared in Table 3 to the real retrieval from the simple, weighted and "full" averages a posteriori combinations, and to the NN retrieval in the IR+MW configuration. The synergy is also estimated in Table 3 for the real retrieval (not for the theoretical estimates).

[64] First, it can be noted that the theoretical estimates (for the simple, weighted and full averages) are an underestimation of the real retrieval uncertainties: Instead of 2.8, 2.7 and 2.6 for the theoretical combinations, the actual retrievals have respectively a retrieval uncertainty of 3.6, 3.4 and 3.3 $kg.m^{-2}$ for *TCWV*. Similarly for the retrieval of TCO_3, the actual retrieval uncertainties at the levels of 6.3, 6.1 and 6.0 DU for the theoretical uncertainties are significantly greater than the theoretical estimates (5.5, 5.1 and 4.9). These significant differences prove that the assumptions to estimate the theoretical uncertainties are too simplistic. First, the uncertainties are more complex than what the Gaussian hypothesis states. Second, the independence of the two retrieval errors is not satisfied: Even if the two retrievals are from two different instruments and wavelengths (IR and MW), these uncertainties can be state-dependent which can introduce correlations among them. It is not a surprise that the weighted average is better than the simple averaging because the uncertainty characterization for each source of information (even if it is not perfect) is taken into account: The weighted average will emphasize the observations with lower uncertainties. The full covariance matrix is the best a posteriori combination because it combines more a priori information, namely the correlation of the errors.

[65] Another striking comment is that the a posteriori combination can degrade the best independent retrieval. This can be observed for the *TCWV* retrieval, with negative synergy factors observed (−6.64% for the simple average and −1.88% for the weighted average). The simplistic mixture of the two independent retrievals can degrade the best one if the hypotheses are not correct, as already mentioned (Gaussian character and independence of the two retrieval errors). This could indicate a wrong individual uncertainty assessment, but the errors are well characterized in this example. This shows that the a posteriori combination is a too simplistic approach. For example, the individual retrieval uncertainties are dependent on the state of the atmosphere, so the weighting of the two independent retrievals should use such state-dependency.

[66] The comparison with the NN retrieval from the IR + MW configuration clearly shows that using simultaneously the two sources of information within the retrieval has much

better performance than just combining a posteriori the individual retrievals. The synergy is much better exploited with the a priori data fusion of the satellite data principle. The error decreases from 3.3 to 2.9 for the *TWWV* when using the NN instead of the a posteriori combination, which represents about 15% decrease of the retrieval errors. For the TCO_3, the error decreases from 6.0 to 4.9 which represents a 20% decrease of the retrieval errors. Both improvements are very significant and justify the use of the NN approach to better use synergy of the satellite observations. This should not be surprising [*Aires*, 2011]: Using all the sources of information simultaneously allows addition of information, if the summation is done optimally, with reliable assumptions (i.e., additive synergy), reduce the uncertainties when the information is redundant (i.e., de-noising synergy), but also, and this cannot be done with a posteriori combination, exploit the interaction terms (i.e., non-linear synergy). These non-linear interactions between the various input satellite observations make it possible to account for the advantages and deficiencies of each satellite observation. For example, when a satellite observation is saturated for a particular range of the variable to retrieve, the other observations can help the retrieval. The NN data fusion is able to coherently combine the two sources of information (IR from IASI and MW from AMSU-A + MHS) in a way that depends upon the atmospheric situation.

6. Conclusion and Perspectives

6.1. Conclusions

[67] A retrieval chain has been designed to retrieve the atmospheric profiles from the MetOp-A satellite, exploiting the synergy between different measurements available from this operational platform. This satellite provides coincident observations in the IR (from IASI) and in the MW (from AMSU-A and MHS), with nadir geometries. This work focused on the major atmospheric parameters, namely temperature, water vapor and ozone profiles, for which the selected MetOp-A instruments are particularly sensitive. This work focused on clear-sky situations over ocean. Although real observations are used in the analysis, the results that are obtained are dependent on the training data set that is used to calibrate the neural network model. In particular, the correlations of ozone with temperature and water vapor can be questioned considering the quality of ozone in the ECMWF analysis. This is an important factor since these correlations are used in the indirect synergy for the retrieval of ozone.

[68] The developed method is very general and flexible and can be adapted to other applications, i.e., other variables, instruments, or environmental conditions. The synergy measures proposed in this study use a NN retrieval scheme, but any retrieval algorithm could be used instead.

[69] The results obtained here with real satellite observations confirm the theoretical results derived from simulations [*Aires et al.*, 2011b]. The major conclusions from this study are as follows.

[70] 1. The NN approach is very efficient to exploit the synergy due to its truly multivariate nature and its non-linear capacities (not all retrieval methodologies can benefit from the synergy between observations, as shown in *Aires et al.* [2011b]).

[71] 2. Strong synergies exist between the microwave and the infrared for the retrieval of atmospheric temperature and water vapor.

[72] 3. Simple statistical retrieval tools can realistically measure the potential synergy of a set of satellite observations, this is particularly interesting in the development phase of the satellite missions.

[73] 4. The synergetic data fusion/merging of the satellite observations in the retrieval scheme is more efficient than the a posteriori combination of products from independent retrievals.

[74] Although the retrieval of atmospheric profiles under clear-sky conditions over ocean is already considered of reasonable quality when using one type of instrument only, the synergetic merging of the observations of different instrument improves the results. This study proves that there is still potential improvement in the retrieval of key atmospheric variables such as temperature, water vapor, or ozone profiles if synergy is used, even for supposedly "easy" conditions. The efficient use of simultaneous observations in various wavelength ranges makes it necessary to develop new retrieval strategies, as presented here. The variational assimilation developed in numerical weather prediction centers also benefits from the instrument synergy but if the goal is to obtain pure satellite data sets, to validate global circulation models for instance, methods have to be implemented to use the synergy among all available satellite observations. The NN approach proved its efficiency in this framework.

6.2. Perspectives

[75] The topic of synergy exploitation is very attractive to the satellite remote sensing community since many existing retrieval methods were designed for retrieving single atmospheric or surface variable by using dedicated sensor channel(s). Although there are many causes that may lead the breakdown of a retrieving method, the utilization of partial or limited information is by no means likely to build up the "panorama" of an object. A synergistic method that combines different sources of information about an identical object might provide the way to the full image of the object. With the advent of more and more hyper-spectral and high-resolution satellite sensors, synergistic retrieving methods may find more and more applications.

[76] The potential use of the synergy between the different observations opens exciting perspectives: The more difficult the problem is, the higher the potential benefits of synergy are. The visible information from the GOME II instrument is used to retrieve ozone [*Burrows et al.*, 2005; *Loyola et al.*, 2011]. It has been shown that IASI provides ozone information as well [*Aires et al.*, 2002b; *Coheur et al.*, 2005]. Therefore, it would be very interesting to test the ozone retrieval using the synergy among the IR and visible observations.

[77] In this study, all observations were performed from the same platform. Note that the synergy between observations can also be applied to instruments on different platforms, although its practical application can be less convenient due to the necessity to have adequate spatial and temporal matching. The Sentinel suite (ESA missions for the GMES (Global Monitoring for the Environment and Security) program) would certainly benefit from the synergetic

use of different observations from different satellites. The characteristics of the spatiotemporal mismatches errors could be introduced into the neural network retrieval scheme. The focus was here on passive observations, but synergy can also be found between passive and active measurements at similar frequencies. This is the concept behind a mission such as Soil Moisture Active and Passive (SMAP) [*Piles et al.*, 2004].

[78] In this first study, only ocean cases have been considered. The next natural step is to analyze the potential of using simultaneously IR and MW observations over the continents. In the MW as well as in the IR, several factors contribute to make retrieval of atmospheric profiles much more difficult over land than over ocean. First, surface temperatures and emissivities are much more variable in space and time over land than over ocean. Second, in the microwave, the land surface emissivities are much higher than the ocean ones, making the surface contribution to the signal much larger. Finally, land surface emissivities are very complex to model, from arid surfaces to dense vegetation or snow, being dependent upon a large number of surface parameters that are difficult to estimate, on a global basis (e.g., soil moisture, soil roughness, lithology, snow cover and properties).

[79] Several efforts have been conducted to develop global data sets of land surface emissivities at both microwave and infrared frequencies [*Prigent et al.*, 2006; *Seemann et al.*, 2008; *Zhou et al.*, 2011], directly calculated from satellite observations. The use of a priori MW emissivity information has shown, in research mode, its potential to improve the retrieval of atmospheric parameters (i.e., temperature and water vapor) over land [*Aires et al.*, 2001; *Karbou et al.*, 2005; *Aires et al.*, 2011b]. Some work has also been done in the infrared [*Seemann et al.*, 2008]. A particularly interesting idea would be to combine the MW and the IR observations for the retrieval of atmospheric profiles over continents. Using all wavelength ranges will improve significantly the characterization of the surface, and the retrieval would benefit from a higher constraint in the inversion process, especially in the lower atmosphere. It is suggested here to extend the use of MW and IR observations from MetOp-A, for the retrieval of atmospheric profiles over land, with the help of the tools recently developed to estimate the land surface emissivities in the MW and in the IR.

[80] Remote sensing under cloudy conditions would also certainly benefit from the synergy between the VIS, IR, and MW domains. First, multi-wavelength observations will benefit the retrieval of cloud characteristics [*Aires et al.*, 2011a]. Second, with clouds better constrained, the atmospheric retrieval will be facilitated. MW measurements are much less sensitive to clouds than the VIS and the IR, and to some extent, they can provide information in the clouds and below. Thin cirrus are essentially transparent at MW frequencies up to 200 GHz. Liquid clouds mostly interact with the MW radiation through emission/absorption and their effect can be accounted for in the retrieval of the atmospheric temperature and water vapor profiling. Convective clouds with a significant ice phase can scatter the MW radiation and their effect will be more difficult to take into account, likely limiting the accuracy of the profile retrieval. As a consequence, convective and precipitating situations should be avoided in a first attempt to evaluate the synergy of the VIS, IR and MW observations for atmospheric profiling. Under cloudy non-precipitating conditions, the MW, IR, and VIS measurements offer complementary information about the clouds. The IR provides its top height and its temperature. Preliminary information on the optical thickness of clouds can be derived from the VIS. Using simultaneously these different types of measurements will better constrain the inversion process and the retrieval will benefit from it. The inversion methodology developed for clear sky conditions over ocean can be adapted to the cloudy cases.

[81] The synergy evidenced in this study is very significant and should be taken into account in the design of the instruments for the new missions. The instrument characteristics should be determined not separately, independently for each sensor but instead all the instruments should be taken into account, to optimize globally the whole observing system. The tools we developed could be adopted to simulate the effect of the different potential channel characteristics and combinations to reach an optimum set of channels across wavelength ranges. This will not only ensure optimal retrieval accuracy, but also cost efficiency for the system, avoiding any non-necessary redundancies from an instrument to the other.

[82] **Acknowledgments.** This project has been funded by ESA General Studies Programme (GSP) under contract "Toward a synergetic approach for the retrieval of atmospheric geophysical parameters from optical/infrared and microwave measurements", 21837/08/NL/HE. We would like to express our gratitude to Björn Rommen and Marc Bouvet from ESA/ESTEC for interesting discussions related to this project.

References

Aires, F. (2011), Measure and exploitation of multi-sensor and multi-wavelength synergy for remote sensing: 1. Theoretical considerations, *J. Geophys. Res.*, *116*, D02301, doi:10.1029/2010JD014701.

Aires, F., C. Prigent, W. Rossow, and M. Rothstein (2001), A new neural network approach including first-guess for retrieval of atmospheric water vapor, cloud liquid water path, surface temperature and emissivities over land from satellite microwave observations, *J. Geophys. Res.*, *106*(D14), 14,887–14,907.

Aires, F., W. B. Rossow, N. A. Scott, and A. Chédin (2002a), Remote sensing from the infrared atmospheric sounding interferometer instrument: 1. Compression, denoising, and first-guess retrieval algorithms, *J. Geophys. Res.*, *107*(D22), 4619, doi:10.1029/2001JD000955.

Aires, F., W. B. Rossow, N. A. Scott, and A. Chédin (2002b), Remote sensing from the infrared atmospheric sounding interferometer instrument: 2. Simultaneous retrieval of temperature, water vapor, and ozone atmospheric profiles, *J. Geophys. Res.*, *107*(D22), 4620, doi:10.1029/2001JD001591.

Aires, F., F. Bernardo, H. Brogniez, and C. Prigent (2010), An innovative calibration method for the inversion of satellite observations, *J. Appl. Meteorol. Climatol.*, *49*(12), 2458–2473.

Aires, F., F. Marquisseau, C. Prigent, and G. Sèze (2011a), A land and ocean microwave cloud classification algorithm derived from AMSU-A and -B, trained using MSG-SEVIRI infrared and visible observations, *Mon. Weather Rev.*, *139*, 2347–2366.

Aires, F., M. Paul, C. Prigent, B. Rommen, and M. Bouvet (2011b), Measure and exploitation of multi-sensor and multi-wavelength synergy for remote sensing: 2. An application for the retrival of atmospheric temperature and water vapour from MetOp, *J. Geophys. Res.*, *116*, D02302, doi:10.1029/2010JD014702.

Bishop, C. M. (1996), *Neural Networks for Pattern Recognition*, 482 pp., Clarendon, Oxford, U. K.

Bowman, K., et al. (2006), Tropospheric emission spectrometer: Retrieval method and error analysis, *IEEE Trans. Geosci. Remote Sens.*, *44*(5), 1297–1307.

Brenner, N., S. P. Strong, R. Koberle, and W. Bialek (2000), Synergy in neural code, *Neural Comput.*, *12*, 1531–1552.

Burrows, J., et al. (2005), The Global Ozone Monitoring Experiment (GOME): Mission concept and first scientific results, *J. Atmos. Sci.*, *56*, 151–175.

Chalon, G., F. Cayla, and D. Dievel (2001), IASI: An advanced sounder for operational meteorology, paper presented at 52nd Congress of IAF, IAF, Toulouse, France.

Christi, M. J., and G. L. Stephens (2004), Retrieving profiles of atmospheric CO_2 in clear sky and in the presence of thin cloud using spectroscopy from the near and thermal infrared: A preliminary case study, *J. Geophys. Res.*, *109*, D04316, doi:10.1029/2003JD004058.

Coheur, P., B. Barret, S. Turquety, D. Hurtmans, J. Lazaro, and C. Clerbaux (2005), Retrieval and characterization of ozone vertical profiles from a thermal infrared nadir sounder, *J. Geophys. Res.*, *110*, D24303, doi:10.1029/2005JD005845.

Cybenko, G. (1989), Approximation by superpositions of a sigmoidal function, *Math. Control Signals Syst.*, *2*, 303–314.

Dubovik, O., A. Smirnov, B. Holben, M. King, Y. Kaufman, T. Eck, and I. Slutsker (2000), Accuracy assessments of aerosol optical properties retrieved from Aerosol Robotic Network (AERONET) Sun and sky radiance measurements, *J. Geophys. Res.*, *105*, 9791–9806, doi:10.1029/2000JD900040.

EUMETSAT (2009), ATOVS level 2 product guide, report, EUMETSAT, Darmstadt, Germany. [Available at oiswww.eumetsat.org/webops/eps-pg/arovs-l2/atovsl2-pg-index.htm.]

Garand, L., et al. (2001), Radiance and Jacobian intercomparison of radiative transfer models applied to HIRS and AMSU channels, *J. Geophys. Res.*, *106*, 24,017–24,031.

George, M., et al. (2009), Carbon monoxide distributions from the IASI/MetOp mission: Evaluation with other space-borne remote sensors, *Atmos. Chem. Phys. Discuss.*, *9*, 9793–9822, doi:10.5194/acpd-9-9793-2009.

Hewison, T., and R. Saunders (1996), Measurements of the AMSU-B antenna pattern, *IEEE Trans. Geosci. Remote Sens.*, *34*(2), 405–412.

Holben, B. N., et al. (1998), Aeronet—A federated instrument network and data archive for aerosol, *Remote Sens. Environ.*, *66*, 1–16, doi:10.1016/S0034-4257(98)00031-5.

Hornik, K., M. Stinchcombe, and H. White (1989), Multilayer feedforward networks are universal approximators, *Neural Networks*, *2*, 359–366.

Kalnay, E., et al. (1996), The NCEP/NCAR 40-year reanalysis project, *Bull. Am. Meteorol. Soc.*, *77*, 437–470.

Karbou, F., F. Aires, C. Prigent, and L. Eymard (2005), Potential of Advanced Microwave Sounding Unit-A (AMSU-A) and AMSU-B measurements for atmospheric temperature and humidity profiling over land, *J. Geophys. Res.*, *110*, D07109, doi:10.1029/2004JD005318.

Landgraf, J., and O. P. Hasekamp (2007), Retrieval of tropospheric ozone: The synergistic use of thermal infrared emission and ultraviolet reflectivity measurements from space, *J. Geophys. Res.*, *112*, D08310, doi:10.1029/2006JD008097.

Liu, Y., R. Parinussa, W. Dorigo, R. D. Jeu, W. Wagner, A. van Dijk, M. McCabe, and J. Evans (2011), Developing an improved soil posiure dataset by blending passive and active microwave satellite-based retrievals, *Hydrol. Earth Syst. Sci.*, *15*, 425–436, doi:10.5194/hess-15-425-2011.

Loyola, D., et al. (2011), The GOME-2 total column ozone product: Retrieval algorithm and ground-based validation, *J. Geophys. Res.*, *116*, D07302, doi:10.1029/2010JD014675.

Maddy, E. S., C. Barnet, and A. Gambacorta (2009), A computationally efficient retrieval algorithm for hyperspectral sounders incorporating a priori information, *IEEE Geosci. Remote Sens. Lett.*, *6*(4), 802–806, doi:10.1109/LGRS.2009.2025780.

Mo, T. (1996), Prelaunch calibration of the advanced microwave sounding unit-A for NOAA-K, *IEEE Trans. Microwave Theory Tech.*, *44*, 1460–1469.

Piles, M., D. Entekhabi, and A. Camps (2004), A change detection algorithm for retrieving high-resolution soil moisture from SMAP radar and radiometer observations, *IEEE Trans. Geosci. Remote Sens.*, *47*(12), 4125–4131.

Pommier, M., et al. (2010), IASI carbon monoxide validation over the Arctic during POLARCAT spring and summer campaigns, *Atmos. Chem. Phys. Discuss.*, *10*(10), 14,445–14,494, doi:10.5194/acpd-10-14445-2010.

Prigent, C., F. Aires, and W. Rossow (2006), Land surface microwave emissivities over the globe for a decade, *Bull. Am. Meteorol. Soc.*, *87*, 1573–1584, doi:10.1175/BAMS-87-11-1573.

Razavi, A., C. Clerbaux, C. Wespes, L. Clarisse, D. Hurtmans, S. Payan, C. Camy-Peyret, and P. Coheur (2009), Characterization of methane retrievals from the IASI space-borne sounder, *Atmos. Chem. Phys.*, *9*, 7889–7899.

Rodgers, C. (1990), Characterizatioon and error analysis of profiles retrieved from remote sounding measurements, *J. Geophys. Res.*, *95*(D5), 5587–5595.

Rumelhart, D., G. Hinton, and R. Williams (1986), Learning internal representations by error propagation, in *Parallel Distributed Processing*, edited by D. E. Rumelhart and J. L. McClelland, pp. 318–362, MIT Press, Cambridge, Mass.

Seemann, S. W., E. E. Borbas, R. O. Knuteson, G. R. Stephenson, and H.-L. Huang (2008), Development of a global infrared land surface emissivity database for application to clear sky sounding retrievals from multispectral satellite radiance measurements, *J. Appl. Meteorol. Climatol.*, *47*, 108–123.

Shannon, C. E. (1948), A mathematical theory of communications, *Bell Syst. Tech. J.*, *27*, 379–423.

Smirnov, A., B. N. Holben, T. F. Eck, O. Dubovik, and I. Slutsker (2000), Cloud-screening and quality control algorithms for the aeronet database, *Remote Sens. Environ.*, *73*, 337–349, doi:10.1016/S0034-4257(00)00109-7.

Uppala, S., et al. (2005), The ERA-40 re-analysis, *Q. J. R. Meteorol. Soc.*, *131*, 2961–3012.

Vaughan, G., C. Cambridge, L. Dean, and A. Phillips (2005), Water vapour and ozone profiles in the midlatitude upper troposphere, *Atmos. Chem. Phys.*, *5*, 963–971.

Worden, H., M. Deeter, D. Edwards, J. Gille, J. Drummond, and P. Nédélec (2010), Observations of near-surface carbon monoxide from space using MOPITT multispectral retrievals, *J. Geophys. Res.*, *115*, D18314, doi:10.1029/2010JD014242.

Worden, J., X. Liu, K. Bowman, K. Chance, R. Beer, A. Eldering, M. Gunson, and H. Worden (2007), Improved tropospheric ozone profile retrievals using OMI and TES radiances, *Geophys. Res. Lett.*, *34*, L01809, doi:10.1029/2006GL027806.

Zhou, D. K., A. M. Larar, X. Liu, W. L. Smith, L. L. Strow, P. Yang, P. Schlüssel, and X. Calbet (2011), Global land surface emissivity retrieved from satellite ultraspectral IR measurements, *IEEE Trans. Geosci. Remote Sens.*, *49*, 1277–1290, doi:10..1109/TGRS.2010.2051036.